attribute or feature (e.g. presence or absence, location, correlation, identity, etc.) or a molecule, cell, tissue or plant. As used herein, a marker nucleic acid is a nucleic acid molecule that is utilized to determine an attribute or feature (e.g., presence or absence, location, correlation, etc.) or a molecule, cell, tissue or plant.

[0229] Nucleic acid molecules of the present invention can be used to monitor expression. A microarray-based method for high-throughput monitoring of gene expression may be utilized to measure gene-specific hybridization targets. This 'chip'-based approach involves using microarrays of nucleic acid molecules as gene-specific hybridization targets to quantitatively measure expression of the corresponding genes (Schena et al., *Science* 270: 467-470 (1995); Shalon, Ph.D. Thesis, Stanford University (1996)). Every nucleotide in a large sequence can be queried at the same time. Hybridization can be used to efficiently analyze nucleotide sequences.

[0230] Several methods have been described for fabricating microarrays of nucleic acid molecules and using such microarrays in detecting nucleotide sequences. For instance, microarrays can be fabricated by spotting nucleic acid molecules, e.g. genes, oligonucleotides, etc., onto substrates or fabricating oligonucleotide sequences in situ on a substrate. Spotted or fabricated nucleic acid molecules can be applied in a high density matrix pattern of up to about 30 non-identical nucleic acid molecules per square centimeter or higher, e.g. up to about 100 or even 1000 per square centimeter. Useful substrates for arrays include nylon, glass and silicon. See, for instance, U.S. Pat. Nos. 5,202,231; 5,445,934; 5,525,464; 5,700,637; 5,744,305; 5,800,992. Sequences can be efficiently analyzed by hybridization to a large set of oligonucleotides or cDNA molecules representing a large portion of genes of a genome. An array consisting of oligonucleotides or cDNA molecules complementary to subsequences of a target sequence can be used to determine the identity of a target sequence, measure its amount, and detect differences between the target and a reference sequence. Nucleic acid molecule microarrays may also be screened with molecules or fragments thereof to determine nucleic acid molecules that specifically bind molecules or fragments thereof.

[0231] A nucleic acid "library" encompassed by the present invention also includes biochemical libraries of the nucleic acids of SEQ ID NOS:1-9112, e.g., collections of nucleic acids representing the provided nucleic acids. The biochemical libraries can take a variety of forms, e.g. a solution of cDNAs, a pattern of probe nucleic acids stably bound to a surface of a solid support (microarray) and the like. By array is meant an article of manufacture that has a solid support or substrate with one or more nucleic acid targets on one of its surfaces, where the number of distinct nucleic acid sequences may be in the tens, or even hundreds, or even thousand, or even tens of thousands. Each nucleic acid will comprise at 18 consecutive covalently linked nucleotides, and often at least 25 consecutive covalently linked nucleotides, and often at least 100 to 1000 consecutive covalently linked nucleotides, and may represent up to a complete coding sequence or cDNA. A variety of different array formats have been developed and are known to those of skill in the art. The arrays of the subject invention find use in a variety of applications, including gene expression analysis, drug screening, mutation analysis and the like, as disclosed in the above-listed exemplary patent documents.

[0232] In addition to the above nucleic acid libraries, analogous libraries of polypeptides are also provided, where the where the polypeptides of the library will represent at least a portion of the polypeptides encoded by SEQ ID NOS:1-9112.

[0233] The microarray approach may also be used with polypeptide targets (U.S. Pat. No. 5,445,934; U.S. Pat. No. 5,143,854; U.S. Pat. No. 5,079,600; U.S. Pat. No. 4,923,901). Essentially, polypeptides are synthesized on a substrate (microarray) and these polypeptides can be screened with either protein molecules or fragments thereof or nucleic acid molecules in order to screen for either protein molecules or fragments thereof or nucleic acid molecules that specifically bind the target polypeptides (Fodor et al., *Science* 251: 767-773 (1991)).

[0234] It is understood that one or more of the molecules of the present invention, preferably one or more of the nucleic acid molecules or protein molecules or fragments thereof of the present invention may be utilized in a microarray based method. In a preferred embodiment of the present invention, one or more of the *D. v. virgifera* nucleic acid molecules or protein molecules or fragments thereof of the present invention may be utilized in a microarray based method. A particular preferred microarray embodiment of the present invention is a microarray comprising nucleic acid molecules encoding genes or fragments thereof that are homologues of known genes or nucleic acid molecules that comprise genes or fragments thereof that elicit only limited or no matches to known genes. A further preferred microarray embodiment of the present invention is a microarray comprising nucleic acid molecules having genes or fragments thereof that are homologues of known genes and nucleic acid molecules that comprise genes or fragment thereof that elicit only limited or no matches to known genes.

[0235] In a preferred embodiment, the microarray of the present invention comprises at least 10 nucleic acid molecules that specifically hybridize under high stringency to at least 10 nucleic acid molecules encoding *D. v. virgifera* protein or fragments thereof of the present invention. In a more preferred embodiment, the microarray of the present invention comprises at least 100 nucleic acid molecules that specifically hybridize under high stringency to at least 100 nucleic acid molecules that encode a *D. v. virgifera* protein or fragment thereof of the present invention. In an even more preferred embodiment, the microarray of the present invention comprises at least 1,000 nucleic acid molecules that specifically hybridize under high stringency to at least 1,000 nucleic acid molecules that encode a *D. v. virgifera* protein or fragment thereof of the present invention. In a further even more preferred embodiment, the microarray of the present invention comprises at least 2,500 nucleic acid molecules that specifically hybridize under high stringency to at least 2,500 nucleic acid molecules that encode a *D. v. virgifera* protein or fragment thereof of the present invention. In a preferred embodiment, at least 50%, preferably at least 70%, more preferably at least 80%, even more preferably at least 90% of the nucleic acid molecules that comprise the microarray contain one protein or fragment thereof.

[0236] Nucleic acid molecules of the present invention may be used in site directed mutagenesis. Site-directed mutagenesis may be utilized to modify nucleotide sequences, particularly as it is a technique that allows one or more of the amino acids encoded by a nucleic acid molecule to be altered (e.g. a threonine to be replaced by a methionine). Three basic methods for site-directed mutagenesis are often employed. These are cassette mutagenesis (Wells et al., *Gene* 34: 315-23 (1985)), primer extension (Gilliam et al., *Gene* 12: 129-137